Wessel van der Loo 7 **Florence Mougel**
María S. Sánchez 7 **Christianne Bouton** 7 **Enrico Castien**
Amélia Fonseca 7 **Nuno Ferrand** 7 **Ramón Soriguer**
Monique Monnerot

Cytonuclear disequilibria in wild populations of rabbit (Oryctolagus cuniculus L.) suggest unequal allele turnover rates at the b locus (IGKC1)

Abstract DNA sequence comparisons suggest that evolutionary rates at the rabbit *IGKC1* locus can differ among allelic lineages. Here we address the question of whether population *turnover* rates can vary among *IGKC1* alleles. We studied the distribution of sixteen *IGKC1* (or *b*-locus) allotypes in areas comprising the aboriginal species range (Iberian peninsula). Rabbits in this area belong to one of two distantly related mitochondrial lineages (mtDNA types) *A* and *B*. In the more recent distribution area of the species, all rabbits belong to the mtDNA type *B* lineage, and *IGKC1* alleles *b4* and *b5* comprise over 90% of the gene pool. These two alleles are also predominant in areas of mtDNA type *B* prevalence within the Iberian range. However, in areas of mtDNA type *A* prevalence, the b4 and b5 allotypes are rare or absent; they apparently have been replaced by serologically related, but distinct, 'endemic' variants. The cytonuclear disequilibria were highly significant, also within the subsample consisting of populations from Spain. These observations suggest that allelic persistence times for the predominant *IGKC1* lineages could be shorter than the divergence time of the major mtDNA lineages *A* and *B*. In contrast, the relative gene frequencies of the *IGKC1* allele *b9* were similar among the type *A* and type *B* rabbits; it was present in most populations at low frequency. In consequence, persistence times of the *b9* allele appear to be longer than the divergence time of lineages *A* and *B*. The data reported here are in agreement with the DNA sequence data, providing further proof that the molecular clock can run at different rates among allelic lineages at the rabbit *IGKC1* locus.

Key words Cytonuclear disequilibria 7 Diversity enhancement selection 7 Molecular clock 7 Histocompatibility antigens 7 Immunoglobulin

W. van der Loo (Y)
Institute of Molecular Biology, Evolutionary ImmunoGenetics Research Group, Vrije Universiteit Brussel, 65 Paardestraat, B-1640 St-Genesius Rode, Belgium
e-mail: wvdloo@ben.vub.ac.be, Tel.: c32-235-90245, Fax: c32-235-90390

C. Bouton
Vlaams Interuniversitair Instituut voor Biotechnologie; Parasitology and Ultrastructure, VUB, B-1640 St-Genesius Rode, Belgium

F. Mougel 7 M. Monnerot
Centre de Génétique Moléculaire, CNRS, F-91198 Gif-sur-Yvette Cedex, France

M.S. Sánchez [1]
Museo Nacional de Ciencias Naturales, Departemento de Biologia Evolutiva, 2 José Gutierrez Abascal, E-28006 Madrid, Spain

A. Fonseca
Departamento de Biologia, Universidade dos Açores, PT-9502 Ponta Delgada, Portugal

Enrico Castien
Servicio Conservacion de la Naturaleza, Gobierno de Navarra E-31002, Pamplona, Spain

N. Ferrand
Centro de Estudos em Ciência Animal, ICETA, Universidade do Porto, P-4480 Vila do Condo, Portugal

R. Soriguer
CSIC, Estacion Biologica de Doñana, E-41013, Sevilla, Spain

[1] *Present address*:
Department of Genetics, The University of Georgia, Athens, GA 30602, USA

Introduction

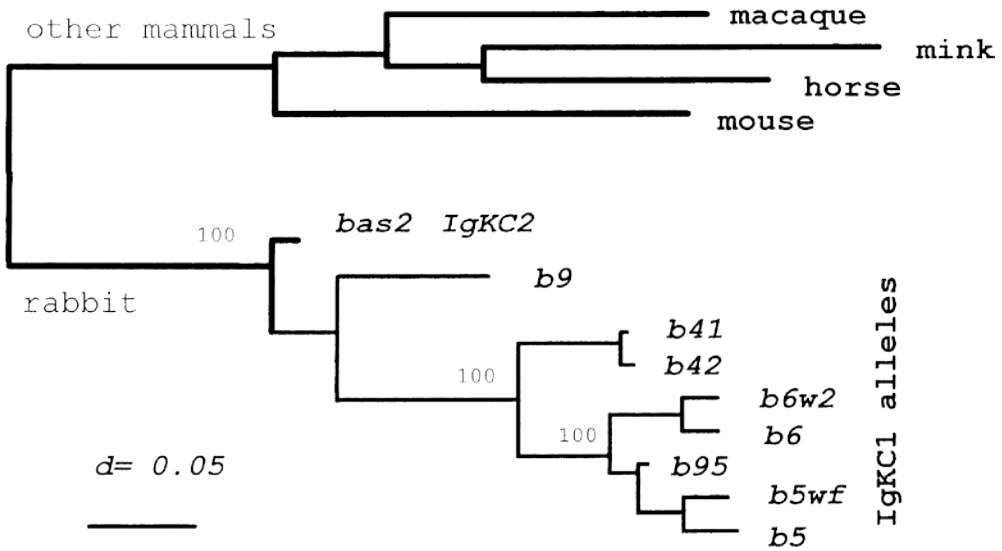
The molecular and population genetic processes leading to the high allelic diversity observed at protein loci involved in mechanisms of recognition and defense are primary issues in evolutionary genetics. Evidence of diversifying selection has been documented mainly for genes at the major histocompatibility complex (MHC) loci; Hughes and Nei 1989; Klein 1986), olfactory and pheromone receptors (Merrit et al. 1998; Ngai et al.

1993), and immunoglobulins (Tanaka and Nei 1989; van der Loo 1987, 1993). Antibody-encoding genes (*Ig* genes) are of special interest because they do not exist as such in the germline: these genes are generated during cell differentiation by recombination of gene segments that are scattered along the chromosome (Tonegawa 1983). In the rabbit species, some *Ig* gene fragments are used (or rearranged) more often, compared with isotypic counterparts. For the variable regions of the heavy (H) chain, preferential usage of the *IGHV1* gene is well documented (Becker and Knight 1990; Short et al. 1991; for the here adopted IMGT (ImMunoGeneTics) nomenclature see *http://imgt.cnusc.fr: 8104*). The expression of light (L) chain genes of the kappa class (K) also shows a very strong bias in gene choice. Although two functional isotypic gene segments are available for the constant region (*IGKC1* and *IGKC2*, respectively), the vast majority of rabbit antibodies feature the K1 L-chain isotype which is characterized by the IGKC1 constant domain (Benamar and Cazenave 1982; Garcia et al. 1982; Heidmann and Rougeon 1983). Both the predominantly expressed *IGHV1-* (alias *a-*) and *IGKC1-* (alias *b-*) locus gene display an unusually high level of allelic diversity, with alleles at the *IGKC1* locus differing at up to 41% amino acid (aa) positions (Emorine et al. 1984; Mage et al. 1987; Reisfeld et al. 1965). The serological expression of this allelic variation was referred to as Allotypie (Oudin 1960). In contrast, genetic variation at the quasi-silent *IGKC2* gene (*bas* locus) is similar to that at classical protein loci, where alleles differ by 1–2% aa substitutions (Mariamé et al. 1987). DNA sequence comparisons of rabbit *IGKC1* alleles revealed that substitution rates were significantly larger at aa replacement sites than at synonymous sites (Bernstein et al. 1983a; Emorine et al. 1984; van der Loo and Verdoodt 1992), while noncoding regions were more conserved than exon regions (Akimenko et al. 1986). The hypothesis that this unusual form of allelic diversity is due to forces favoring (antigenic) variation at the protein level was furthermore supported by population genetic studies which strongly suggested overdominance-type selection is occurring at the *IGKC1* locus (van der Loo 1993).

These observations are indicative of a possible relationship between the level of gene usage and of (adaptive) gene diversity. Interestingly, in rabbits that are heterozygous at these loci *(IGHV1* and *IGKC1*), one allele can be expressed more than the other. This phenomenon, which is called *allelic imbalance in gene usage* or *pecking order*, appears to be a constitutive characteristic of the allotype (Akimenko et al. 1986) and is related to allelic exclusion. For example, in a heterozygous *IGKC1-b4/b9* rabbit, lymphocytes expressing the b4 allotype are about four times higher in number than lymphocytes committed to b9 expression (Lummus et al. 1967; Mage 1967; Mage et al. 1981). This creates a unique situation where relationships between gene usage and evolutionary patterns can be assessed for lineages at the same locus. Here we raise the question of whether the predominantly expressed alleles at the rabbit *IGKC1* locus (e.g., the *b4* allele) evolve at a higher rate compared with those alleles that are underexpressed when present in heterozygous condition (e.g., the *b9* allele).

One line of evidence can be drawn from sequence comparisons. The phylogenetic tree in Fig. 1 is representative of trees obtained with a variety of programs applied to the DNA or protein sequence alignments of mammalin *IGKC* exons. This tree highlights the singu-

Fig. 1 Best phylogenetic tree of mammalian *IGKC* coding regions. Trees were constructed using the program *dnaml* of the PHYLIP package version 3.572c (*dnaml*, J**P**25,15; default settings) applied to the DNA sequence alignment which corresponds to that of amino acid sequences show in Fig. 6. Bootstrapping values 1 95% are indicated at the major nodes of the rabbit genes. *bas2*: rabbit *IGKC2* gene; *b9-b5*: eight rabbit *IGKC1* alleles. The molecular clock hypothesis is not supported, even when the analysis is restricted to the seven rabbit *IGKC1* alleles *b9, b42, b95, b6, b6w2, b5, b5wf.* Indeed, $x^2_6 \geq 2^*$ {Ln Likelihood $_{dnamlk}$ - Ln Likelihood$_{dnaml}$}**P**21.5; $P \sim 0.005$



lar evolution of the leporid IGKC coding regions, and indicates that evolutionary rates might be (1) lower for the *IGKC2* gene (*bas*) than for *IGKC1* alleles (*b9-b5 branch*), and (2) lower for the *b9* allele than for alleles of the *b4-b5* branch. Furthermore, maximum likelihood analyses of DNA sequences (Felsenstein 1990) strongly suggest that the molecular clock can run at different rates among lineages of the rabbit *IGKC1* locus (cf. legend Fig. 1). However, because genic exchange among rabbit *IGKC* genes can not be excluded or is even likely (Ayadi 1991), sequence comparisons alone do not allow us to reject the generally accepted hypothesis of monotonous *evolutionary* rates of alleles at the same locus.

In this paper, we verify the population genetic fundament of the above hypothesis: namely, that *turnover rates* of alleles at the same locus are monotonous. To best delimit the turnover rates at *IGKC1* loci, we analyzed the variation in allele frequencies among two distantly related maternal lineages. We will present evidence indicating that turnover rates might be substantially slower in the lineage of the underexpressed b9 allotype than in the lineages of the predominantly expressed b4, b5, and b6 allotypes.

The European rabbit (*Oryctolagus cuniculus*) originated in the Iberian Peninsula, where there are two morphologically and genetically differentiated subspecies, *O. c. cuniculus* and *O. c. algirus* (Lopez-Martinez 1977). Restriction data on whole mitochondrial DNA (mtDNA) revealed two major maternal lineages, *A* and *B* (Biju-Duval et al. 1991). Lineage *A* predominates in the range of *O. c. algirus* (southern and south-western areas and the Azorean Islands). Lineage *B* prevails in the northern and eastern parts of the peninsula. Sequence comparisons of mtDNA revealed a 4% nucleotide (nt) divergence between the two lineages, suggesting more than 1 million years (My) of separate evolution. Domestic breeds and wild rabbits of continental Europe (north of the Pyrenean Mountains), Great Britain, and overseas belong to the subspecies *O. c. cuniculus* and display the mtDNA types of the *B* lineage (Hardy et al. 1995; Monnerot et al. 1994). These areas will here be referred to as the *recent* area of distribution or RAD (details on distribution and history of the species are found in Flux 1994; Callou 1995).

Four *IGKC1* alleles are currently distinguished in the RAD: *b4, b5, b6,* and *b9*. They are serologically identical to those occurring in domestic breeds (allotypes) and, for convenience, are called *domestic* alloleles. With a mean frequency above 60%, the *b4* allele was prevalent at the vast majority of investigated localities from RAD, and more than 90% of the gene pool consisted of just two alleles, *b4* and *b5* (Cazenave et al. 1987; Herd and Edmonds 1977; van der Loo 1987, 1993). This frequency hierarchy [p_{b4} 1 (p_{b5} ⊂ p_{b6}) 1 p_{b9}] was also observed in domestic breeds, and its close match with the pecking order in gene expression (*b4* 1 *b5* ⪰ *b6* ⫐ *b9)* provides a first indication of a possible causal relationship between allelic imbalance in gene usage and population parameters that are liable to affect evolutionary rates.

We extended these studies to populations of the aboriginal range of the genus. In a previous study on wild populations of Andalusia and Portugal, the four *domestic IGKC1* alleles were reported to occur at roughly similar frequencies (each around 10%), together with at least seven *endemic* variants (van der Loo et al. 1991). We have now reassessed these data in the light of recent studies on the distribution of mtDNA lineages in the Ibero-Lusitanian region (including Azores). Furthermore, rabbits were collected in areas of (1) mtDNA type *B* prevalence (Navarra), (2) mtDNA type *A* prevalence (Andalusia), and (3) (apparent) introgression of both lineages (Estremadura). For these sample populations, individual rabbits were analyzed for mtDNA as well as for sixteen serologically defined *IGKC1* alleles (*b*-locus allotypes). The frequency distribution of these alleles and their association with the major maternal lineages were estimated.

Materials and methods

Collection sites

All specimens were from wild populations. Spain: *Navarra* (*NV*): five localities south of the city of Tafalla; *Estremadura* (*EM*): Badajoz; *Andalusia* (*AN*): Las Lomas (Cadiz area). Continental Portugal: *Santarem* (*ST*), *Vila-Viçosa* (*VV*). Azores (Atlantic Ocean, Portugal): *Flores* (*FL*). France: seven localities (*FR*). Belgium and Netherlands: three localities (*BNL*). Great Britain: six localities (*GB*). Australia: nine localities (*AUS*). All rabbits were more than 3 months old. Tissue and/or serum samples were stored and shipped frozen. Sites of the Iberian range are localized in Fig. 2.

Mitochondrial DNA typing

Two fragments of mitochondrial DNA were used for typing. They were obtained by the polymerase chain reaction (PCR) using the following primers:

Cytb3: 5b-ATGAAACTGGCTCCAACAAC-3b
Th3: 5b-CTCCATTCCGGCTACAAGAC-3b
Pro1 5b-CCACCTCAGCACCCAAAGCT-3b
NC4: 5b-ATGGCCCTGAGGGAAGACC-3b
NC5: 5b-CTTAATAAAACTCAAGTACTTC-3b

Fragment (1) is a 586 base pair (bp) PCR product representing the 3b part of the *cytochrome* gene, and is obtained with primers *Cytb3* and *Th3*. The digestion with *Hin* fI and *Bgl* II allows the distinction of mtDNA types *A* vs *B*, and *Alu* I sites ensure the discrimination of *B3* and *B4* among *B* types. Fragment (2) corresponds to 565 bp of the 5b part of the noncoding region, as amplified using primers *Pro1* and *NC4*; the PCR products were partially sequenced using an internal primer *NC5*, and were digested with restriction enzyme *Rsa* I. The restriction fragment length analysis allows the discrimination of type *B1*, while sequencing helps the assignment of types *B8* to *B11* and their subtypes (Biju-Duval et al. 1991; Mougel 1997). The mtDNA types and subtypes were determined for the rabbits from Spain. Subtypes were not distinguished for rabbits from Portugal and the Azorean islands. The mtDNA types of rabbits of the RAD (*FR, BNL, GB,* and *AUS*) were determined only for a relatively small number of representatives.

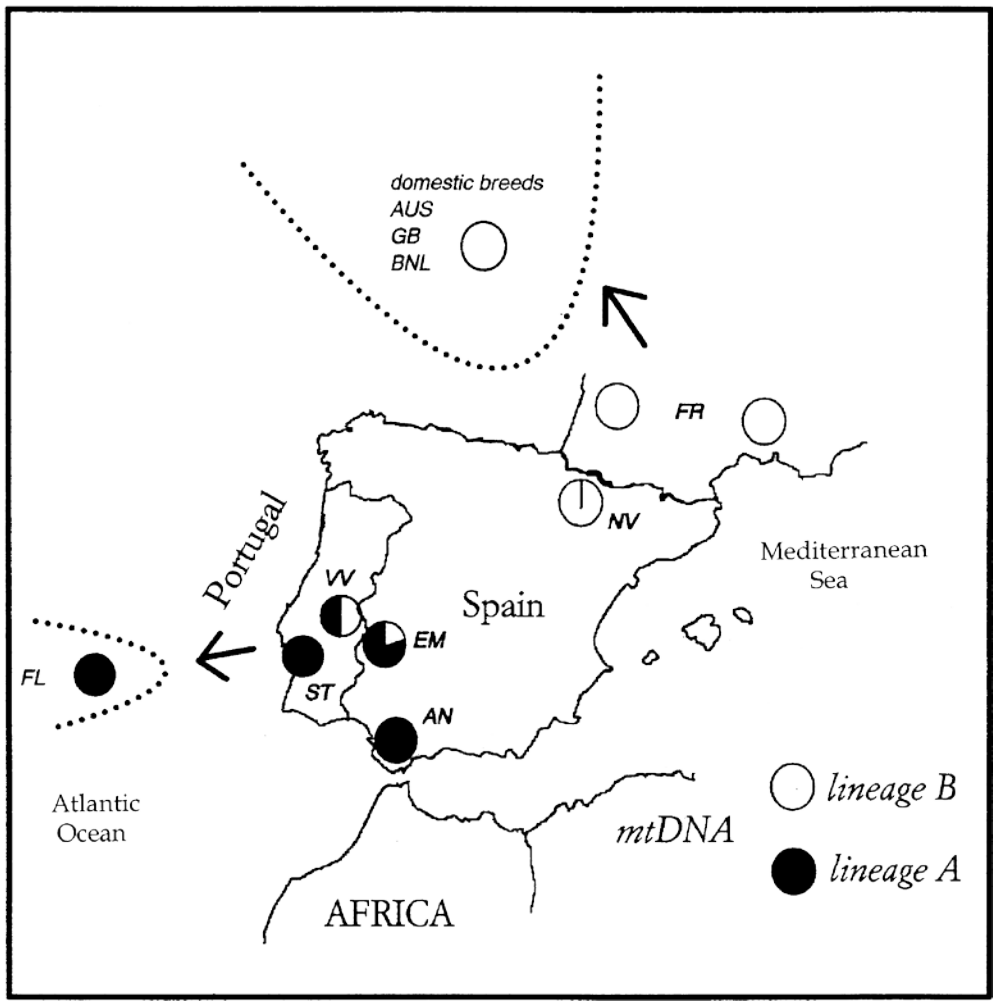


Fig. 2 Sample location and mtDNA-type frequency distribution of European rabbit: Spain: *NV*: Navarra; *EM*: Estremadura; *AN*: Andalusia. Portugal: *VV*: Vila-Viçosa; *ST*: Santarem; *FL*: Island of Flores (Azorean archipelago). *FR*, *BNL*, *GB*, and *AUS* refer to France, Belgium-and-Netherlands, Great Britain, and Australia, respectively. Wild rabbits of Flores originate from Portugal and belong exclusively to mtDNA lineage *A*. Domestic rabbits and wild rabbits of continental Europe, Great Britain, and Australia originate from southern France and/or eastern Spain and belong to mtDNA lineage *B*

IGKC1 typing

The serological typing was carried out as in van der Loo and co-workers (1991), by *immunodiffusion*, using a panel consisting of four allo-antisera (aAs) and three monoclonal antibodies (mAb). The aAs were raised in domestic rabbits following Kelus and Gell (1967); mAb were produced in mouse as outlined in van der Loo and co-workers (1983). Sixteen *IGKC1* allotypes were distinguished by serology in crossreacting reference systems as shown in Fig. 3. Genotypes were inferred from the phenotypic arrays (Table 1; cf. van der Loo et al. 1991,1995,1998). Sera with phenotypes that could not readily be resolved into their putative allelic components were subjected to double gel immunodiffusion (Kelus and Steinberg 1991), a technique which allows one to verify whether determinants recognized by two different antibodies are epitopes of the same or of different molecular species in the test serum. All sera were furthermore tested for the two alleles (*e14* and *e15*) of the *e* locus, as described in van der Loo and co-workers (1996). This locus encodes the *CH2* domain of the Ig gamma heavy chain constant region (Bernstein et al. 1983b).

Endemic allotypes nomenclature

We adopt here a working nomenclature which includes the name of the *domestic* allotype to which the serological distance appears smallest and a symbol separated by lowercase *w* for workshop. The symbol refers to the type of cross-reactivity by which the allotype was defined. So far, sequence data were in perfect agreement with the serological inferences (van der Loo et al. 1995, 1998). The smallest interallelic protein distance observed until now between the here distinguished allotypes was 0.092 (i.e., *b5* vs *b5wf*; cf. van der Loo et al. 1998). Note that two *b4* lineages *b41* and *b42* (or *b4* and *b4var*) exist in domestic breeds that differ by two aa substitutions (Sogn and Kindt 1976). The reagents used in this study do not permit one to distinguish these subtypes.

Supertypes

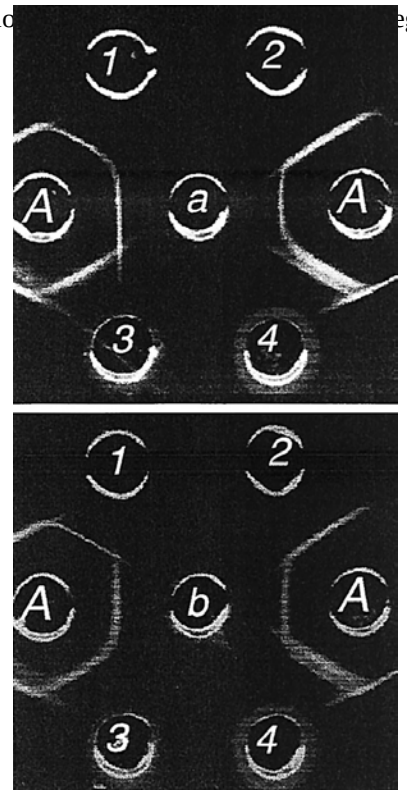
The concept *supertype* is used to partition the numerous alleles at *MHC* loci into a small number of subsets according to serologic relationships (reviewed in Sidney 1996). Lineages that are characterized by b4-specific epitopes were temptatively assigned to the *b4** supertype; of these, those that are *endemic* to the Ibero-Lusitanian region are designated as *b4*ibl* lineages; likewise for lineages *b5* and *b6*. The aAs used do not detect crossreactivity between the *b5* and *b6* allotypes, because b5-specific aAs were produced in heterozygous *b4/b6* rabbits and b6-specific aAs in *b5/b4* rabbits. However, the serological and structural similarity of allotypes b5 and b6 is well documented (Dreher et al. 1990, cf. Fig. 1), and would justify classifying them into one supertype (b5-b6*). We did not observe sera showing cross-reaction with *b9* allotype-specific aAs or mAb.

Estimation of gene diversity

Parameters of gene diversity were estimated according to Nei (1973) and Chakraborty (1974). The gene diversity in the total sample at locus A with alleles A_i was estimated as

$$H_T = 1 - \sum_i p_{Ai}^2,$$

lar evolutio egions, and gene usage and population parameters that are liable to



In the example shown, four sera of individual rabbits were numbered *1–4* and confronted in two different protocols with the same b4-specific antiserum (aAs 5186) disposed in *wells* labeled "A". *Top:* centre well contains the b4wa reference serum (a). *Bottom*: centre well contains the reference for allotype b5wb (b). Sera *1* and *4* show the motif of the b4wa allotype (epitopes *a* in the *b4 column* in Table 1); sera *3* and *4* show b4-associated determinants of the b5wb and b5wf allotype (*b* in the *b4 column* in Table 1). Serum *2* displays determinants of the b4wd allotype (reference not shown). Serum *4* was confirmed as heterozygous *b5wb/b4wa* by further assays

and that within n subdivisions as

$H_S = \sum S_k (1 - \sum S_i p_{Aik}^2)/n,$

where p_{Ai} is the mean frequency of the A_i allele, and p_{Aik} is p_{Ai} at locality k.

The degree of genetic differentiation between populations was estimated by

$F_{ST} = 1 - H_S/H_T.$

The analysis was unweighted given the large differences in sample size. An unbiased weighted estimate of interlocality variance, u, was obtained following Weir and Cockerham (1989).

Cytonuclear disequilibria

Associations of an allele *A* at a nuclear locus (nucleotype *A*) with a gene type *M* at a cytoplasmic locus (cytotype *M*) were measured

Table 1 Serologic characteristics of IGKC1 allotypes

Allotype			Allo-antiserum (aAs)				Monoclonal antibody (mAb)		
			#5186	#5128	#5260	#5837	#2.15	#24C2	#21G10
			raised against allotype:						
	RAD[a]	IBL[a]	b4	b5	b6	b9	b4	b5%b6	b6
b4	C	C	**1**	0	0	0	**1**	0	0
b4we	P	C	**e**	0	0	0	**1**	0	0
b4wd	P	C	**d**	z	z	0	0	**1**	0
b4wc	P	C	**c**	z	0	0	0	0	0
b4wa	P	C	**a**	0	0	z	0	0	0
b5	C	C	0	**1**	0	0	0	**1**	0
b5ws	P	C	0	**s**	0	0	0	**1**	0
b5we	P	C	0	**e**	0	0	0	**1**	0
b5wd	P	C	0	**d**	0	0	0	**1**	0
b5w0b	P	C	0	**b**	0	0	0	**1**	0
b5wf	P	C	**b**	**f**	0	0	0	**1**	0
b5wb	P	C	**b**	**b**	0	0	0	**1**	0
b6	C	C	0	0	**1**	0	0	**1**	**1**
b6w2	P	C	0	0	**1**	0	0	0	**1**
b6w5	P	C	0	0	**1**	0	0	**1**	0
b9	C	C	0	0	0	**1**	0	0	0

Reactions were recorded as type '1' in the case of a perfect fusion of the precipitation lines of test serum and reference serum of the nominal allotype (identity reactions). Negative reactions are recorded as (0). Cross-reactions were recorded as (a–z); 'z' refers to (weak) precipitation reactions that were not used for typing. The bold symbols (**a–s**) refer to clear, strong cross-reactions that were identified by a particular reference serum/antiserum system (see Fig. 3). Allotypes showing such strong reactions with aAs#5186 belong to the *b4** supertype, those reacting with aAs#5128 to the *b5** supertype (b5wf and b5wb belong to both supertypes). Data concerning the *domestic* allotypes are shaded throughout Tables 1–4

[a] "C" observed, "P" not observed in RAD or in IBL; "RAD": recent area of distribution (France, Belgium, Netherlands, Great Britain, Australia); "IBL": Ibero-Lusitanian range (Spain, continental Portugal, Azores)

by the cytonuclear disequilibrium (Asmussen and Arnold 1991):

$D_{AM} = p_{AM} - p_A * p_M$,

where p_A and p_M are the frequencies of the nucleotype *A* and cytotype *M*, respectively, and p_{AM} is the observed frequency of nucleotype *A*-cytotype *M* associations (cytonucleo-type *AM*). By defining $D_{AM}{}^{max}$ as the theoretical maximum value of D_{AM} for a sample with gene frequencies p_A and p_M, a standardized expression of D_{AM} is given by Z_{AM},

$Z_{AM} = D_{AM}/D_{AM}{}^{max}$,

where

$D_{AM}{}^{max} = p_A(1 - p_M)$ if $p_A \sim p_M$,

$D_{AM}{}^{max} = p_M(1 - p_A)$ if $p_A 1 p_M$.

For subdivided populations, the disequilibrium between genes *A* and *M* estimated for the total sample (D_T) consists of components reflecting the correlation in frequencies p_A and p_M between localities (Cov), and the total disequilibria within localities (D_s), respectively. If p_{Ak} and p_{Mk} are p_A and p_M at locality k, n_k is the sample size at locality k, and $w_k = n_k/S_k n_k$, then:

$D_S = S_k w_k \{p_{AM} - p_{Ak} * p_{Mk}\}$,

$\mathrm{Cov}(A,M) = S_k w_k \{p_{Ak} * p_{Mk} - p_A * p_M\}$.

and by consequence:

$D_T = D_S + \mathrm{Cov}(A,M)$.

We also determined trigenic disequilibria D_{AAM} (Weir and Cockerham 1989) and genotypic disequilibria D^{AAM} and their constraints as defined by Asmussen and Basten (1996). The higher order effects were not significant and always marginal compared with the digenic component $D_{AM,}$ and therefore are not shown.

Computer software

Use was made of the computer facilities offered by the Belgian EMBnet Node *(http://www.be.embnet.be)*. The software programs used were those provided by the EGCG extension of the WISCONSIN PACKAGE 8.1.0 of the University of Wisconsin and by the PHYLIP PACKAGE, version 3.572c by J. Felsenstein (University of Washington).

Results

The serological analysis allows one to distinguish clearly 16 different IGKC1 allotypes in the total sample. The procedures used may underestimate the number of alleles, but an over-estimate of gene diversity is very unlikely (see van der Loo et al. 1991). Estimates of the relative frequencies of these alleles and heterozygosity levels are shown for each locality in Table 2. A detailed study on gene correlations in the RAD was presented in van der Loo (1993). For the populations from Spain (*AN*, *EM*, and *NV*) and Flores (*FL*), exact Hardy-Weinberg distribution tests were carried out. Departures from expected homozygosity levels did not exceed 10% and were not significant, except on occasion when expected numbers were $\sim$1 (not shown). Allele frequencies in the remaining samples were estimated assuming Hardy-Weinberg equilibrium. For one Portuguese sample (*VV*), we had to rely on data obtained in a previous study, where different types of cross-reacting epitopes were not identified by reference sera. Heterozygosity levels were very high (H_S in Table 2), and were highest (1 80%) in populations where both mtDNA lineages were present. *IGKC1* gene diversity was higher among mtDNA type *A* rabbits ($H_{T/A}$) than among type *B* rabbits ($H_{T/B}$; Table 3). None of the rabbits collected outside the RAD expressed the e14 allotype, while all displayed identity reactions with e15-specific aAs (indicating *e15* homozygosity in the Ibero-Lusitanian range).

Interlocality variation of IGKC1 alleles

The vast majority of rabbits were found to display determinants belonging to the b4 or b5 motifs; that is to say, most rabbits harbor genes of the *b4** or of the *b5** supertype. The parameters F_{ST} displayed in Table 2 measures the degree of frequency variation (or genetic differentiation) for a particular allele across localities or areas. Interlocality variation was clearly higher for the *b4* than for the *b9* allele. This was confirmed by the unbiased estimates of within-locality correlations of gene pairs, u according to Weir and Cockerham (1989; not shown). Interlocality variation was much less pronounced for the *b4** and *b5** supertypes when compared to that for the alleles composing these supertypes. However, since differences in F_{ST} or in u values among alleles can not be evaluated statistically (Ewens and Feldmann 1976), the meaning of this observation is merely indicative. Nevertheless, we believe that these marked differences are noteworthy (1) because F_{ST} values are related to allele turnover rates, and (2) because of the similitude with diversity patterns of *MHC* supertypes in human populations: Whereas *HLA-I* supertype frequencies were similar among ethnic groups, subtypes showed large inter-ethnic frequency variations (Sidney 1996).

Cytonuclear disequilibria

Figure 4 highlights the association of *domestic* allotypes with the maternal lineage *B*. This association was very pronounced for the b4 and b5 allotypes (Tabl. 2). By contrast, the frequency distribution of the b9 allotype was not correlated with a mtDNA lineage. Cytonuclear disequilibria were estimated more in detail for the samples collected in Spain (*AN*, *EM*, and *NV*). Every rabbit in this set of samples was characterized both for its mtDNA type and for its *Ig* allotype. In Spain numerous subtypes of the *B* lineage have been distinguished (*B1*-*B11*), of which only a limited number occur in RAD (mostly *B1* to *B4*). A detailed report on the mtDNA types at the localities under study can be found in Mougel (1997). In the present study only two categories of the mtDNA type *B* subtypes were distinguished: one comprising subtypes currently observed in domestic breeds as well as in a wild population of the RAD (*B1*-*B4*, here *Bd*), and one comprising those that are en-

Table 2 Distribution of IGKC1 allotypes in wild populations of rabbit

Allele	Frequency (%) at locality:										F_{ST}[a]
	FL	*AN*	*SN*	*EM*	*VV*[b]	*NV*	*FR*	*BNL*	*GB*	*AUS*	
mtDNA types:											
– **A**	**100**	**100**	**100**	**82**	**52**	**1.5**	–	–	–	–	
– **B**	–	–	–	**18**	**48**	**98.5**	**100**	**100**	**100**	**100**	
IGKC1 allotypes:											
– b4we	32.3	–	–	–	~13.6[d]	–	–	–	–	–	0.30
– b4wd	26.8	26.5	–	6.3	~13.6	–	–	–	–	–	0.16
– b4wc	–	2.9	–	–	~13.6	5.6	–	–	–	–	0.04
– b4wa	–	7.8	–	12.5	~13.6	1.4	–	–	–	–	0.08
– b5wf	–	5.9	–	18.8	~13.5[e]	–	–	–	–	–	0.14
– b5wb	–	35.3	34.1	15.6	~13.5	8.3	–	–	–	–	0.20
S b4*ibl	59.1	78.4	34.1	53.2	27.1	15.3					**0.34**
– b4	–	2.9	18.2	12.5	13.5	48.6	60.8	59	65.5	71.2	**0.31**
S *b4**supertype	**59.1**	**81.3**	**52.3**	**65.7**	**40.6**	**63.9**	**60.8**	**59**	**65.5**	**71.2**	**0.04**
– b5ws	–	1	–	3.1	–	–	–	–	–	–	0.02
– b5wd	–	–	–	–	~8.2[f]	4.2	–	–	–	–	0.04
– b5w0b	40.9	–	27.3	–	~8.2	–	–	–	–	–	0.27
– b5we	–	2.9	–	–	~8.2	2.1	–	–	–	–	0.02
– b5wf	–	5.9	–	18.8	~13.5[e]	–	–	–	–	–	0.14
– b5wb	–	35.3	34.1	15.6	~13.5	8.3	–	–	–	–	0.20
S b5*ibl	40.9	45.1	61.4	37.5	21.7	14.6	–	–	–	–	**0.28**
– b5	–	–	–	9.4	4.2	4.9	29.5	39	25.1	26.L	**0.16**
S *b5**supertype	**40.9**	**45.1**	**61.4**	**46.9**	**25.9**	**19.5**	**29.5**	**39**	**25.1**	**26.7**	**0.07**
– b6w2	–	6.9	6.8	3.1	~30.2[g]	16.7	–	–	–	–	0.08
– b6w5	–	–	3.4	6.3	~30.2	–	–	–	–	–	0.04
– b6	–	1	3.4	3.1	~30.2	6.9	3.3	–	–	–	0.06
S *b6**supertype	–	**7.9**	**13.6**	**12.5**	**30.2**	**23.6**	**3.3**	**–**	**–**	**–**	**0.13**
– b9	–	**6.9**	**6.8**	**9.4**	**16.8**	**1.4**	**6.4**	**2**	**9.4**	**2**	**0.04**
H_S[c]	0.66	0.78	0.76	0.88	(0.86)	0.68	0.54	0.50	0.51	0.42	
Sample size	52	51	22	16	33	72	656	234	463	1100	

The frequency of b4*ibl allotypes is the sum of frequencies of *endemic* alleles expressing one or more major determinants of the b4 motif (Table 1). *Endemic* are those alleles that do not occur in RAD. The *b4** supertype groups all alleles that express determinants of the b4 motif (*b4* and *b4*ibl*); likewise for *b5* and *b6*. For details see Materials and methods. The *FL-AUS* localities are identified in Fig. 1

[a] F_{ST} estimates the component of gene diversity due to differences between localities (VV not included)

[b] ~indicates that subtypes were not distinguished; the figure corresponds to the sum of gene frequencies of these subtypes: [d]b4wa-b4we; [e]b5wf-b5wb; [f]b5wd-b5we; [g]b6-b6w5

[c] H_S is the heterozygosity at the *IGKC1* locus level within localities (see Materials and methods)

demic in the Iberian range (*Be*). (Note: in the case of the Portuguese sample, the material used for mtDNA typing and for serological typing, although collected at the same site, was obtained from different rabbits).

The relative allele frequencies within rabbits subdivided according to their mtDNA type are shown in Fig. 5. The frequencies of two-gene cytonuclear genotypes (*cytonucleo-types*) in Table 3 show that the more frequent alleles were associated with distinct maternal lineages (i.e., *b4* with lineage *B; b4wd, b5wb,* with lineage *A*). These correlations of *IGKC1* alleles with the major mtDNA lineages were significant (Table 3). They are clearly *not* due to effects *within* population. In the Estremadura sample, where both maternal lineages were present, (only) 10% of the domestic *IGKC1* genes were found among the mtDNA type *B* rabbits, which represented 18% of the sample (not shown). Because $D_T \geq D_S \geq Cov(A,M)$, cytonuclear equilibrium within localities ($D_S \geq 0$) was confirmed by the fact that values of cytonuclear disequilibria were very similar to the interlocality covariances in gene frequencies ($D_T \geq Cov(A,M)$; Table 3). The (lack of) correlations of *IGKC1* alleles with either *domestic* or *endemic* mtDNA *B* subtypes in the Navarra population is evaluated statistically in Table 4.

Discussion

The analysis of the frequency correlations between the different alleles belonging to the rabbit's nuclear *IGKC1* locus and its two mtDNA lineages may provide new insights into the history of the species, and to the mechanisms involved in the maintenance of unusual gene diversity at this antibody locus.

The presence in wild Iberian populations of the four *"domestic"* allotypes together with ten or more en-

Table 3 Cytonuclear disequilibria and gene diversity in Spain (3 localities, 140 rabbits)

Allotype	p_{bA} [a] %	p_{bB} [a] %	$Cov_S(b, B)$	$D_T(b, B)$	Z (b, B) [b] %	X_1^{2} [c]
b4wd	10.00	0.36	0.0688	0.0519	93	32.7
b5wb	13.93	5.00	0.0540	0.0514	50	19.4
b5wf	3.93	0.36	0.0050	0.0194	84	10.3
b9	3.57	0.71	0.0140	0.0158	68	6.9
b4wa	3.93	1.07	0.0153	0.0161	60	6.1
b5ws	0.71	–	0.0032	0.0038	100	2.3
b6w5	0.71	–	0.0025	0.0038	100	2.3
b5we	1.07	1.43	P0.0005	P0.0009	7	0.0
b5	1.07	2.50	P0.0070	P0.0059	35	1.1
b5wd	–	1.79	P0.0108	P0.0064	64	2.2
b4wc	1.07	2.86	P0.0066	P0.0075	41	1.7
b6	0.71	3.57	P0.0123	P0.0128	64	4.5
b6w2	2.86	8.57	P0.0240	P0.0245	46	6.7
b4	2.50	25.36	P0.1035	P0.1043	80	61.0
Total	46.43	53.57				
Diversity [d]	$H_{T/A}$	$H_{T/B}$	H_T			
	83.4	**73.0**	**85.1**			

Disequilibria (D_T) are very pronounced for alleles occurring at frequencies above 5% and are almost entirely due to frequency correlations of *IGKC1* alleles and mtDNA types between localities (D_T P D_S C Cov_S)

[a] Fraction of genes in the total sample consisting of *IGKC1* alleles associated with mtDNA type *A* or *B*, respectively (i.e., cytonucleo-type frequencies)

[b] D/D_{max}

[c] Chi-square values testing the hypothesis of cytonuclear equilibrium ($D_T(b, B)$; 0)

[d] *IGKC1*-locus heterozygosity within cytotype classes ($H_{T/A}$, $H_{T/B}$) and in the total sample (H_T)

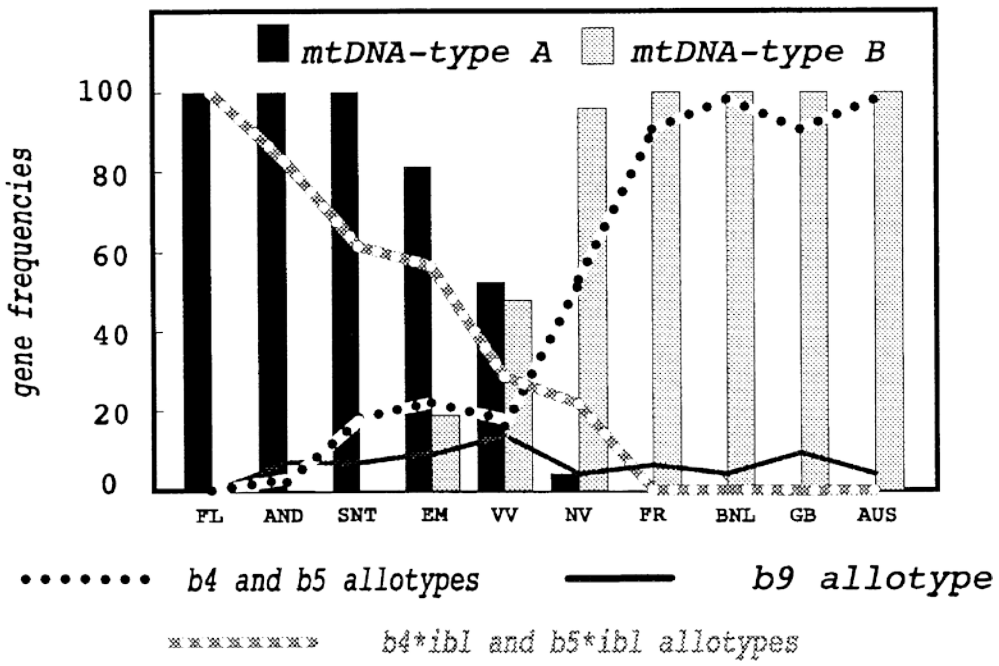


Fig. 4 Frequency correlations between *IGKC1* alleles and mtDNA types: The pronounced genetic differentiation according to mtDNA type coincides with marked differences in gene frequencies of the *b4* and *b5* alleles. In areas of mtDNA type *A* prevalence, the *b4* allele appears to be replaced by *endemic* alleles presenting b4 motifs (i.e., *b4*ibl* allotypes; likewise for the *b5* allele. By contrast, the distribution of the *b9* allele is not correlated to one or the other mtDNA lineage. Geographic localities are indentified in Fig. 2. Localities *FL* through *NV* are part of the Ibero-Lusitanian area (*IBL*); *FR* through *AUS* are part of the recent distribution range (*RAD*)

demic variants is in agreement with previous reports (Cazenave et al. 1987; van der Loo et al. 1991). The existence of a frequency cline of the domestic types which parallels the North-South cline of mtDNA type *B* has not yet been documented. In the Navarra sample more than 60% of the gene pool consisted of domestic allotypes. In Estremadura, this was 34%, in Andalusia 10%, and 0% on Flores. The pronounced geographical cline is due to the predominant domestic allotypes *b4* and *b5*, which, taken together, show a very strong association with the mtDNA type *B* (Figs. 4–5, Tables 2, 3).

Within localities, the frequency distribution of *IGKC1* alleles tends to be hierarchical, except for the sample of Estremadura and Vila-Viçosa, where both maternal types coexist. In Andalusia, with eleven allotypes, more than 60% of the gene pool consisted of just two alleles (*b4wd* and *b5wb*). In Navarra, with ten

Table 4 Cytonuclear correlations within Navarra (72 rabbits)

b Allele	$p_{b/Be}$[a] %	$p_{b/Bd}$[b] %	$p_{b/Bd}$[c] % ! 3	$D_t(b, Bd)$	Z (*b, Bd*) %	X_1^{2}[d]
b4wc	5.56	–	–	P0.0139	100	2.8
b5wb	7.64	0.69	2.07	P0.0139	66	1.9
b9	1.39	–	–	P0.0035	100	0.7
b4wa	1.39	–	–	P0.0035	100	0.7
b6w2	**13.19**	3.47	**10.41**	P0.0069	16	**0.3**
b5	3.47	1.39	4.17	0.0017	4	0.1
b5wd	2.78	1.39	4.17	0.0035	11	0.2
b4	**35.42**	13.19	**39.57**	0.0104	2	**0.3**
b6	4.17	2.78	8.34	0.0104	20	1.3
b5we	–	2.08	6.24	0.0156	100	9.2
Domestic[e]	**44.45**	17.36	**52.08**	**0.0191**	**4**	**1.2**
Total	75.00	25.00	75.00			

For most alleles, correlations are not significant because of small sample size. However, for the predominant IGKC1 allotype b4, and for the total of domestic types, the *absence* of disequilibrium seems well established

[a] *IGKC1* cytonucleo-type frequencies for mtDNA types *not* occurring in RAD (type *Be* and *A*)

[b] *IGKC1* cytonucleo-type frequencies for mtDNA types occurring in RAD (type *Bd*)

[c] Latter frequencies adjusted to sample size of *Be* sample

[d] Chi-square values testing the hypothesis $D_T(b, Bd)$P0

[e] Sum of domestic allotypes: b4, b5, b6, and b9

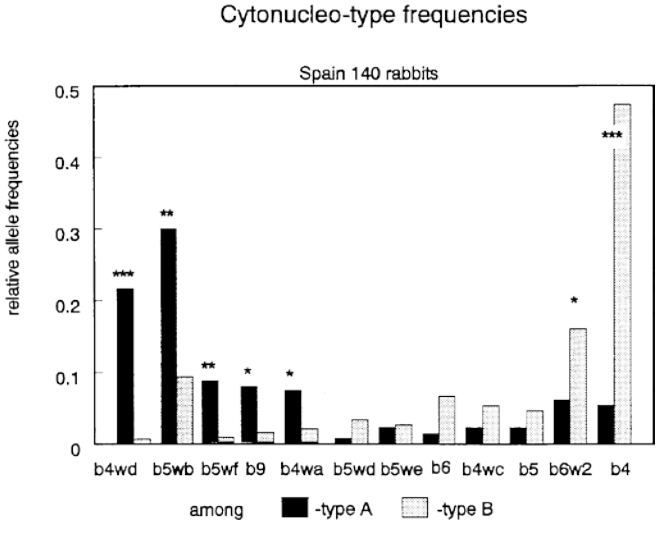


Fig. 5 Relative *IGKC1* allele frequencies in mtDNA lineages *A* and *B*, respectively: Frequencies are shown for the samples collected in Spain (*NV, EM,* and, *AN*). Alleles that are most frequent in one mitochondrial lineage tend to be rare in the other. The probability *P* of cytonuclear equilibrium is indicated for the different alleles: *** $P \sim 0.0001$; ** $P \sim 0.001$; * $P \sim 0.01$

IGKC1 alleles, a single allele (*b4*) represents 50% of the gene pool.

In areas of mtDNA type *B* prevalence, *b4* and *b5* are by far the most frequent alleles. In contrast, in areas of mtDNA type *A* prevalence, the most frequent alleles are endemic representatives of the *b4** and *b5** supertypes (*b4*ibl* and *b5*ibl* alleles). These endemic variants occur at frequencies above 50% in southwestern areas, reaching 100% in the Azorean Island of Flores. It is as if in the latter areas the *domestic b4* – and to a lesser extent the *b5* alleles – are being replaced by alleles belonging to the same superlineage (Fig. 4).

mtDNA types as historical markers

The reduced gene diversity on Flores and in RAD is best explained by founder effects in populations that may have originated from Portugal and from Northern Spain, respectively (cf. Ferrand 1995). The hypothesis that the *domestic* gene types occurring in Northern Spain were introduced during historical times with rabbits of trans-pyrenean origin is indeed unlikely, if only because of the total absence of certain genetic markers that are widespread in domestic rabbits and in wild populations of RAD. One of them, the *e14* marker, was found in all but one of 40 localities of RAD at a mean frequency of 0.21 (van der Loo et al. 1987, 1993, 1996), but was entirely missing in the Ibero-Lusitanian range. Analysis of nuclear satellite DNA polymorphisms has furthermore revealed the existence of alleles that are common in RAD, while absent in the Navarra sample under study (Vachot 1997).

The pronounced frequency correlations between *IGKC1* alleles and mtDNA lineages suggest that the greater diversity in Estremadura and Vila-Viçosa might be due to admixture between two genetically differentiated lineages. Another possibility is that the maternal *A* and *B* lineages and *IGKC1* alleles first differentiated within interbreeding populations: Estremadura and Vila-Viçosa would be remnant representatives of this ancestral population. Afterwards, maternal lineage *B* became extinct in southwestern regions, in concert with the *b4* and *b5* alleles. On the other hand, in northeastern Spain, lineage *A* almost became extinct, while the *b4* allele became the most common allele. In this scen-

ario, the strong fluctuations among *IGKC1* alleles can be explained by local extinction and recolonisation, possibly related to epizootics. It seems, however, unlikely that these processes could account for the observed geographic divide at the level of the mtDNA genotypes, which, furthermore overlaps with the distribution of two subspecies.

Coalescence or admixture?

The fact that *IGKC1* alleles and mtDNA types were at equilibrium *within* populations provides additional evidence of the historical origin of the cytonuclear correlations. If the observed mitochondrial genetic differentiation is indeed the outcome of long lasting geographic isolation (1–2 My), the existence of mixed populations (example Estremadura) implies admixture of maternal lineages in more recent times. Sporadic presence of *b4* genes in areas of mtDNA type *A* prevalence would therefore not necessarily imply that this allele was already present in the common founders of lineages *A* and *B* (coalescence). Field studies indicate that gene exchange between rabbit populations is mainly the outcome of male-specific dispersion (van der Loo et al. 1996; Webb et al. 1995). In consequence, *nuclear* genes should be less affected by the mechanisms underlying the relative genetic isolation between the maternal lineages. This is particularly true for nuclear loci that are exposed to diversity enhancement selection (Asmussen et al. 1989).

Given its high frequency in the *B* lineage, the presence of the *b4* allele in mtDNA type *A* rabbits (Fig. 4) is therefore not only a possible, but a predictable effect of male-mediated gene flow across the boundaries of maternal lineages (admixture). In areas of mtDNA type *A* prevalence, gene flow bringing the *b4* allele into wild populations could also originate from local farms, since, at least in Portugal, the vast majority of farm rabbits were found to express the two most common alleles of domestic rabbits, *b4* and *b5* (van der Loo et al. 1991; N. Ferrand and co-workers, unpublished).

Nevertheless, the possibility that the occasional presence of *b4* genes in the mtDNA lineage *A* is due to coalescence can not be excluded entirely. DNA sequence determination may allow us to settle this important point. If the sharing of allotypes between mtDNA lineages is due to coalescence, then minor sequence differences between these genes are expected, at least in the noncoding regions, in view of their estimated 1–2 My of separation. Identical sequences, on the contrary, would suggest a more recent admixture.

Relevance to IGKC1-locus evolutionary patterns

The mtDNA data permit us to study the evolutionary genetics at the *IGKC1* locus within a better defined historical context. This antibody locus appears to be exposed to diversity enhancement (i.e., balancing or overdominance-type selection, for details see van der Loo 1987, 1993; van der Loo and Verdoodt 1992, van der Loo et al. 1987), and displays highly unusual patterns of gene diversity that are very similar to those reported for MHC loci (Hughes and Nei 1989; Klein 1986; Nei and Hughes 1991; Sidney et al. 1996).

The evolutionary genetics of the MHC polymorphism is the subject of intensive research and debate. One essential question regarding diversity enhancing selection is how it affects evolutionary *rates* at the gene loci concerned (Takahata and Nei 1990). MHC geneticists have estimated that mutant recruitment rates at MHC loci are in fact fairly slow (0.1–0.5% nt substitutions per My). According to Klein and co-workers, selection at MHC loci causes increased allelic persistence times, but rates of mutant recruitment would *not* be affected. Distances between alleles are therefore believed to be proportional to their time of coalescence (Graser et al. 1996; Klein and O'hUguin 1995; Takahata et al. 1992).

These conclusions were drawn from patterns of trans-species polymorphisms under the assumption of homogenous evolutionary rates among allelic lineages. This assumption is justified because parameters that may cause rate variation, between taxa, between genes or between gene regions, should in general not vary among alleles (i.e., factors such as effective population size, bottlenecks, intrinsic mutation rates, population, and social structure; cf. Gibbs et al. 1998; Hoelzer et al. 1998). However, differential evolutionary patterns among alleles have been reported (Slierendregt et al. 1992; Wayne et al. 1997). As outlined in the introduction, there are strong indications of differential evolution rates at the rabbit *IGKC1* locus. Here we will argue that (1) substitution rates were increased, and (2) this increase was not uniform among alleles.

(1) Evolutionary rates have increased: In Table 5, we display the K_a and K_s values for pairwise comparisons of mammalian *IGKC* coding regions (**F**321 bp). K_a and K_s estimate the distance between two coding sequences at nonsynonymous and at synonymous sites respectively (Li 1993; Li et al. 1984). Among species of different orders, K_a for *IGKC* genes varies between 25 and 34%. Assuming divergence times of 80 My (Benton 1990; Li and Graur 1991), this indicates that at replacement sites, substitution rates (r_a) range between 0.16–0.21% per My (or 1.6–2.1 ! 10^{-9} per site per year). The K_a values between rabbit *IGKC1* alleles of the *b9, b4, b5-b6* lineages varies between 13 (*b4 : b95*) and 26% (*b9 : b5*). Clearly, in the absence of substantial rate increases at the rabbit *IGKC1* locus, allelic lineages at this locus must have persisted during much of the mammalian radiation (at face value: 35–50 My for the *b4 : b5* split; 60–83 My for the *b9 : b4* split). The separation of the lagomorph *IGKC* lineage from that of other mammals should, moreover, predate the onset of mammalian radiation by at least 50 My.

K_s values show more variation; while they can exceed 80% between rodent *vs* nonrodent *IGKC* exons

(r_s 1 0.5%/My), among nonrodents these K_s values average around 50% ($r_s \gtrsim 0.3$%/My). Between higher primates synonymous substitution rates r_s could even be lower: K_s between *human* and *macaque* IGKC sequences is 6%, suggesting rates as low as $r_s \gtrsim 0.12$%/My (assuming 25 My of separate evolution; Gingerich 1984). The K_s values between major rabbit *IGKC1* lineages are notably smaller than the corresponding K_a values: they range between K_s(*b4 : b95*) $\gtrsim$ 7% and K_s(*b9 : b5*) $\gtrsim$ 15%. This means that either synonymous substitution rates have slowed down significantly during allelic divergence or that substitution rates accrued substantially at nonsynonymous sites. We would not exclude such a slow-down. It could possibly be due to interallelic gene conversion, which tends to homogenize the gene pool, at least at sites where variation is *not* adaptive. However, it is conflicting with evidence that K_s and K_a are positively correlated (Alvarez-Valin et al. 1998), particularly at loci exposed to selection (Kreitman and Hudson 1991).

Let us suppose that an increase of r_a was merely the outcome of a release in negative selection at the level of the IGKC1 protein (i.e., non-deterministic). If so, rates at replacement sites would probably not exceed rates at synonymous sites of mammalian IGKC. This would imply that the separation between the *b4* and the *b5* lineages ($K_a \gtrsim 16$%) should date back at least some 16–25 My (for $r_a \gtrsim 0.3$–0.5% My). Between the *b5wf* allele and the *b5* allele ($K_a \gtrsim 5.2$%), divergence times should be larger than 5–8 My. This means that only a fraction of the amino acid differences between these alleles has emerged since the population split that underlies the separation of the two maternal lineages (1–2 My). In consequence, also under this hypothesis, most *IGKC1* alleles separated (long) before the two major mtDNA lineages.

The negation of the hypothesis that the large protein distances among rabbit *IGKC1* alleles are due to a substantial increase in evolutionary rates therefore implies that the correlations of alleles of the *b4** and *b5** supertypes and mitochondrial lineages reflects more recent changes in allele distributions. In that case, alleles were currently lost and replaced (turned over) since the time that the maternal lineages separated (Table 2). But, if long allelic persistence times is the process responsible for the large genetic distances observed between *IGKC1* alleles, we would not expect alleles at this locus to go extinct frequently.

On the other hand, the hypothesis that most of the divergence of the *IGKC1* alleles took place after the maternal split implies amplitudes of rate increase that strongly suggest deterministic processes in favor of protein change. The point is that there are only two ways to explain why a given allele that is prevalent in one population is not found in another population. Either this allelic lineage became locally extinct, or, it has changed since the two populations became genetically isolated. In the case under study, the former explanation conflicts with allelic persistence times that are much longer than the divergence time of the two maternal lineages, while the latter is in conflict with normal evolutionary rates. Both interpretations are, however, consistent with an increase in evolutionary rates.

(2) Evolutionary rates vary among allelic lineages: Some allelic lineages, however, obviously *have* been maintained over very long periods. The b6 allotype occurs in areas of type *A* as well as in areas of type *B* prevalence, often together with the b6w2 allotype. DNA sequence comparisons have shown that the *b6w2* gene derives from the *b6* gene lineage (van der Loo et al. 1995). The *b6* lineage is also present in RAD, implying the coalescence of the *b6* genes of the major mtDNA lineages. In addition, DNA sequences of the *IGKC1* homologue in *Lepus* species (i.e., *bla1* and *bla2*) suggest that allele persistence times can be longer than speciation times (Bouton and van der Loo 1997).

It is striking that the more frequently occurring alleles *b4, b4wd,* and *b5wb* were found in close association with one or the other maternal lineage, whereas the rare *b9* allele was present in both lineages (Figs. 4, 5, Table 3). The *b9* allele occurred at low frequency in wild populations of continental Europe, Great Britain and Australia, a situation which was considered as evidence of frequency-dependent selection (Cazenave et al. 1987; Herd and Edmonds 1974; van der Loo 1993). The protein structures and coding sequences of alleles of the *b4* branch (to which the predominantly and intermediately expressed alleles *b4, b5,* and *b6* belong; cf. Fig. 1) share a number of clearly apomorphic character states (Fig. 6). On the other hand, the poorly expressed b9 allotype shows more similarity to the mammalian consensus and to the IGKC2 (bas) isotype than to any of its allelic counterparts [K_a (*bas:b9*) $\gtrsim$ 12.6%, K_a (*bas:b4-b95*) $\gtrsim$ 17–22%; K_a (*b9 : b4-b95*) $\gtrsim$ 20–26%; Table 5, see also Mage 1987]. Note that genomic analysis by Heidmann and Rougeon (1983) have confirmed the true allelic nature of the b4 and b9 allotypes.

In this context, it is remarkable that the cytonuclear correlations suggest, if anything, that allelic persistence times in the lineage of the more derived *b4* genes are *shorter* compared with those in the more primitive *b9* lineage. This is clearly not what we would expect if the large distances between *IGKC1* alleles were solely the outcome of an increase in allele persistence times, as was proposed for the MHC diversity by Klein and coworkers (1993).

The reported observations are, however, in accordance with the hypothesis that diversity enhancement selection can act more efficiently on alleles that occur at high frequencies in populations, than on rare alleles, which in any case mostly occur in heterozygous individuals. Indeed, a higher gene frequency for a given allele not only implies a higher homozygosity level, but also a larger population size. A higher population frequency implies an increased occurrence of mutation events per unit time in a given allelic lineage. In addition, the apparent positive correlation between homozygosity level and gene usage could reinforce the selective impact of a

mutation in the predominant lineages, once it has occurred. The population data are therefore in agreement with population genetic theory and confirm the inferences obtained by DNA sequence comparisons (Fig. 1). Taken together, the data strongly suggest that evolution rates were substantially higher in the lineage of the preferentially and normally used alleles such as *b4*, *b5*, and *b6* (diversifying evolutionary mode) than in lineages of less expressed alleles such as *b9*. These latter alleles have apparently persisted at low frequencies over very long periods, most probably as a consequence of frequency-dependent selection (conservative mode).

Interdomain bond: cause or effect?

As for MHC, allelic divergence at the *IGKC1* locus could be driven by selection-anticipating mechanisms favoring protein diversity *per se*, such as maternal-fetal interactions (van der Loo 1993). It is possible that this has caused collateral variations in marginal fitness among alleles and among their genotypic combinations, in ways that would promote differential evolutionary patterns among alleles.

It is noteworthy that the rabbit IGK1 L chain has taken an important evolutionary step by generating a disulphide bond between the constant domain and the variable domain. Disulphide bonds between Ig domains of a same polypeptide chain have never been observed for any of the numerous members of the Ig protein family, for any species (Halaby and Mornon 1998). This bond might contribute to the unusual degree of allelic variation (Ayadi et al. 1991; McCartney-Francis 1984; van der Loo et al. 1995), but it is also possible that the emergence of the interdomain bond was primarily a *consequence*, rather than a cause, of the particular evolutionary dynamics at this locus.

In conclusion, the available data suggest that a limited number of allelic *IGKC1* lineages have been maintained over very long times. These persisting lineages may correspond to the serological supertypes *b4**, *b5**, *b6** (or *b5***P***b6**), and *b9**. Within these major lineages turnover rates accrued, but relative rate increases can differ substantially among lineages. The unusual diversity patterns of *IGKC1* alleles are therefore due to forces that favor both long lineage persistence times and an increase in mutant recruitment rates within lineages, which, can vary among lineages

The study of the evolutionary patterns at the *IGKC1* locus may provide new insights into specific parameters affecting evolutionary dynamics at multi-allele loci exposed to diversity enhancement selection. The question of whether alleles evolve at similar rates is a key issue in studies of trans-species polymorphisms. The present study shows that the hypothesis of allelic rate monotony can be falsified. It raises furthermore the question of whether a significant variation in evolution rate

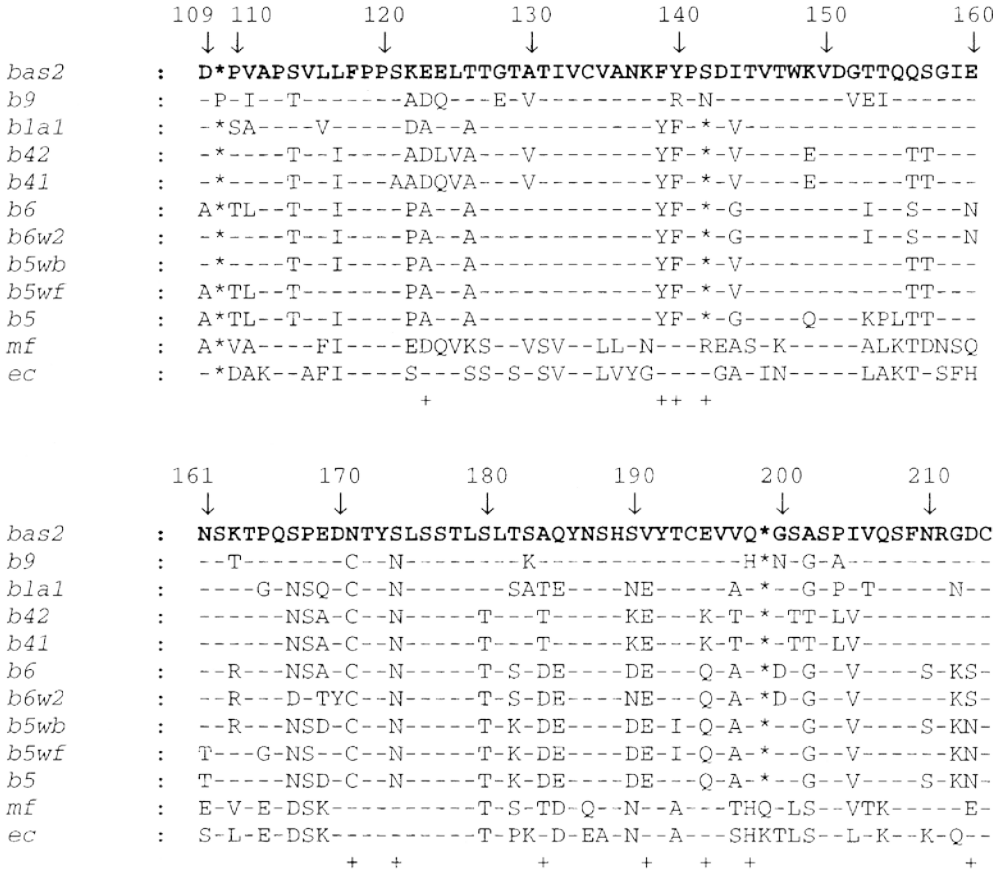


Fig. 6 Alignment of amino acid sequences of IGKC domains. Protein sequences are inferred from the DNA sequences used for the construction of the dendogram in Fig. 1. Numbering of amino acid positions is according to Kabat (1991). b5wf-b9 : *Oryctolagus cuniculus* (IGKC1, different alleles); bla1 : *Lepus americanus* (IGKC1, allele *bla1*); bas2 : *O. cuniculus* (IGKC2, allele *bas2*); mf: *Macaca fascicularis* (macaque IGKC); ec: *Equus caballus* (horse IGKC). *Dashes* indicate identity with bas2; asterisks indicate: indels; a cross indicate positions where bas2 and/or b9 sequences tend to show more similarity to "other" mammals than with to rabbit IGKC1 sequences. GenBank accession numbers are listed in the Table 5 legend

Table 5 Nonsynonymous (K_a) and synonymous substitutions (K_s) at the *IGKC* exon in %

A Among mammalian species [a]

$K_s \backslash K_a$	mm	rr	hs	pt	mf	ss	mv	oa	ec
mm		8.42	**34.40**	**33.15**	**34.19**	**27.84**	**32.86**	**29.28**	**31.73**
rr	19.80		**33.29**	**32.04**	**31.26**	**24.11**	**34.41**	**29.52**	**31.34**
hs	**73.06**	**79.25**		0.75	10.50	**31.16**	**32.84**	**31.61**	**30.26**
pt	**71.51**	**83.60**	2.96		9.56	**29.93**	**31.59**	**30.87**	**29.05**
mf	**74.72**	**76.42**	6.81	5.75		**25.02**	**27.06**	**26.60**	**28.85**
ss	**76.45**	**85.66**	**58.94**	**53.88**	**62.97**		22.58	22.29	27.82
mv	**66.41**	**84.52**	**75.09**	**68.46**	**71.24**	46.25		22.08	29.17
oa	**78.55**	**88.22**	**54.19**	**52.53**	**54.20**	47.02	50.51		29.16
ec	**57.81**	**82.81**	**41.24**	**40.42**	**40.49**	42.83	41.43	34.60	

B Among rabbit *IGKC1* alleles [a]

$K_s \backslash K_a$	bas2	b9	b41	b42	b6	b6w2	b5	b95	b5wf
bas2		2.60	17.57	17.81	21.71	21.53	20.96	18.91	18.15
b9	8.3		**19.81**	**20.06**	**25.88**	**24.30**	**26.40**	**23.75**	**24.49**
b41	17.32	**10.65**		0.99	**17.88**	**17.29**	**16.20**	**13.23**	**14.87**
b42	18.52	**11.69**	0.89		**18.14**	**17.54**	**16.47**	**13.47**	**15.13**
b6	17.14	**12.47**	**8.67**	**9.67**		4.27	7.38	7.27	8.79
b6w2	17.09	**11.32**	**6.63**	**7.62**	1.93		12.10	8.93	12.44
b5	24.03	**14.88**	**10.60**	**11.58**	5.79	5.79		4.97	5.19
b95	21.30	**15.71**	**7.54**	**8.51**	5.87	3.85	4.76		4.76
b5wf	19.69	**13.18**	**9.68**	**10.65**	3.82	5.83	3.80	4.83	

K_a (right above) and K_s (left below) were determined according to Li (1993), applied to the amino acid sequence alignment displayed in part in Fig. 6. Bold figures relate to differences (**A**) between species of different orders (rodents, primates, and ungulates-carnivores); (**B**) between *IGKC1* alleles belonging to branches that are clearly separated according to bootstrap analysis (Fig. 1)

[a] EMBL databank and GenBank accession numbers of the IGKC sequence considered in this study are: (**A**) For mammal species other than leporids: *rr* (rat): J02574; *mm* (mouse): V015613; *hs* (human): M11737; *pt* (chimpanzee): X65287; *mf* (macaque): L13317; *ss* (swine) M59321; *mv* (mink): X75613; *oa* (sheep): X54110; *ec* (horse) X75613. (**B**) For rabbit and hare: *bla1* (snow-shoe hare IGKC1): Z80231; *bas2*: (rabbit IGKC2): M22543; *b5wf-b9* (rabbit *IGKC1* alleles), *b5wf*: AJ003199; *b95*: M22542; *b5*: K01363; *b6*: M37809; *b6w2*: Z48308; *b41*: K01360; *b9*: X00674. An updated rabbit *IGKC* gene table is available in the IMGT (ImMunoGeneTics) database (*http://imgt.cnusc.fr:8104*)

among alleles could exist in the absence of the genic and genetic asymmetries, which prevail at the *IGKC1* locus: allelic exclusion, imbalance of gene expression or hierarchical allele frequencies. For a majority of gene loci, including MHC, these conditions are probably not satisfied. It is fortunate that, compared with many other parameters that influence evolutionary rates, the genic asymmetries mentioned above are less elusive and can be verified by direct observation.

The data reported here provide strong evidence for the coalescence of the *b9*-encoding genes in the common ancestor of the two major maternal lineages *A* and *B*, but argue in favor of shorter coalescence times for the genes belonging to the *b4** and the *b5** supertype. In particular, the evolutionary appearance of the b4 allotype may be more recent than the population split of the mitochondrial lineages. Given that in areas of mtDNA type *B* prevalence (1) the *b4* allele occurs at much higher population frequencies than the *b9* allele, (2) the *b4* allele is preferentially expressed over the *b9* allele in heterozygote individuals, and (3) the structure of the b4 allotype is more derived as compared with that of the b9 allotype, we propose that differences in evolutionary rates among the allelic *IGKC1* lineages contribute to the reported differential cytonuclear associations. This hypothesis can be tested by further sequence determinations and population studies, as well as by a more detailed analysis of the patterns of interallelic nucleotide substitutions.

Acknowledgments This work was financed in part by a grant (Krediet aan Navorsers) allocated to W. van der Loo by the Belgian Founds for Scientific Research (FWO). WvdL is senior research associate at the FWO.

References

Akimenko MA, Mariamé B, Rougeon F (1986) Evolution of immunoglobulin k light chain locus in the rabbit Evidence for differential gene conversion events. Proc Natl Acad Sci USA 83:5180–5183

Alvarez-Valin F, Jabbari K, Bernardi G (1998) Synonymous and nonsynonymous substitutions in mammalian genes: intragenic correlations. J Mol Evol 46:37–44

Asmussen MA, Arnold J (1991) The effects of admixture and population subdivision on cytonuclear disequilibria. Theor Popul Biol 39:273–300

Asmussen MA, Basten CJ (1996) Constraints and normalised measures for cytonuclear disequilibria. Heredity 76:207–214

Asmussen MA, Arnold J, Avise J (1989) The effects of assortive mating and migration on cytonuclear associations in hybrid zones. Genetics 122:923–934

Ayadi H, Marche PN, Cazenave PA (1991) Evolution of the rabbit immunoglobulin kappa chain genes. Immunogenetics 34:201–207

Becker RS, Knight KL (1990) Somatic diversification of immunoglobulin heavy chain VDJ genes: evidence for somatic gene conversion in rabbits. Cell 63:987–992

Benamar A, Cazenave PA (1982) A second rabbit isotype. J Exp Med 156:585–595

Benton MJ (1990) Phylogeny of the major tetrapode groups: morphological data and divergence dates. J Mol Evol 30:409–424

Bernstein KE, Skurla RM Jr, Mage R (1983a) The sequence of rabbit kappa light chains of the b4 and the b5 allotypes differ more in their constant regions than in their 3' untranslated regions. Nucleic Acids Res 11:7205–7214

Bernstein KE, Alexander CB, Mage R (1983b) Nucleotide sequence of a rabbit IgG heavy chain from the recombinant F-1 haplotype. Immunogenetics 18:387–397

Biju-Duval C, Ennafaa H, Dennebouy N, Monnerot M, Mignotte F, Soriguer R, El Gaied A, El Hili A, Mounolou JC (1991) Mitochondrial DNA evolution in Lagomorphs: origin of systematic heteroplasmy organization of diversity in European rabbits. J Mol Evol 33:92–102

Bouton C, van der Loo W (1997) Transspecies nature of the rabbit *b*-locus polymorphism is supported by studies on snowshoe hare (*Lepus americanus*). Immunogenetics 45:444–446

Cazenave PA, Benammar A, Sogn JA, Kindt TJ (1987) Immunoglobulin genes in feral rabbit. In: Dubiski S (ed) The rabbit in contemporary immunological research, 4th Edition, Longman Scientific & Technical, Harlow, pp 148–160

Callou C (1995) Modifications de l'aire de répartition du lapin (*Oryctolagus cuniculus*) en France et en Espagne, du Pléistocène à l'époque actuelle. Etat de la question. Anthropozoologica 21:95–114

Chakraborty R (1974) A note on Nei's measure of gene diversity in a substructured population. Humangenetik 21:85–88

Dreher KL, Asundi V, Wolf B, Bruggeman L (1990) Rabbit Ig *K1b6* gene structure. J Immunol 140:325–330

Emorine L, Sogn JA, Tring FD, Kindt TJ, Max EE (1984) A genomic gene encoding the b5 rabbit immunoglobulin kappa constant region. Proc Natl Acad Sci USA 81:1789–1793

Ewens W J, Feldmann MW (1976) The theoretical assessement of selective neutrality. In: Karlin S, Nevo E (eds) Population genetics and ecology. Academic Press, New York, pp 303–337

Ferrand N (1995) Variaçâo genética de proteínas em populaçôes de coelho (*Oryctolagus cuniculus*). Doctorial thesis, Universidade de Porto; Porto

Flux JEC (1994) World distribution. In: Thompson HV, King CM (eds) The European rabbit: the history and biology of a successful colonizer. Oxford Science Publications, Oxford. pp 8–20

Felsenstein J (1990) PHYLIP manual, Version 3.5 University Herbarium, University of California, Berkeley

Garcia I, Brandt DC, Benammar A, Cazenave PA, Jaton JC (1982) Basilia rabbit express to types of immunoglobulin light chains: lambda and kappa-like. Proc Natl Acad Sci USA 79:4391–4394

Ginderich PD (1984) Primate evolution: Evidence from the fossil record, comparative morphology, and molecular biology. Yrbk Phys Anthrop 27:57–72

Gibbs PEM, Witke WF, Dugaiczyk A (1998) The molecular clock runs at different rates among closely related members of a gene family. J Mol Evol 46:552–561

Graser R, O'hUigin C, Vincek V, Meyer A, Klein J (1996) Transspecies polymorphism of class II *Mhc* loci in danio fishes. Immunogenetics 44:36–48

Halaby DM, Mornon JPE (1998) The immunoglobulin superfamily: an insight on its tissular, species, and functional diversity. J Mol Evol 46:389–400

Hardy C, Callou C, Vigne JD, Casane D, Dennebouy N, Mounolou JC, Monnerot M (1995) Rabbit mitochondrial DNA diversity from prehistoric to modern times. J Mol Evol 40:227–237

Heidmann OL, Rougeon F (1983) Multiplicity of the constant kappa light chain genes in the rabbit genome: a *b4b4* homozygous rabbit contains a *K-bas* gene. EMBO J 2:437–441

Herd ZL, Edmonds JWE (1977) Population genetics of Aa and Ab immunoglobulin allotypes in wild rabbits of South-eastern Australia. I. Allotype frequencies. J Immunogenet 4:315–324

Hoelzer GA, Wallman J, Melnick DJ (1998) The effects of social structure, geographic structure, and population size on the evolution of mitochondrial DNA: II. Molecular clocks and lineage sorting period. J Mol Evol 47:21–31

Hughes A, Nei M (1989) Nucleotide substitution at major histocompatibility complex class II loci: evidence for overdominance selection. Proc Natl Acad Sci USA 86:958–962

Kabat EA, Wu TT, Perry HM, Gottesman KS, Foeller C (1991) Sequences of Proteins of Immunological Interest, US, Department of Health and Human Services National Institute of Health, Bethesda

Kelus A, Gell P (1967) Immunoglobulin allotypes of experimental animals. Prog All 11:141–156

Kelus A, Steinberg CM (1991) Is there a higher rate of mitotic recombination between the loci encoding the VH and CH regions in gonial cells? Immunogenetics 33:255–259

Klein J (1986) Natural History of the major histocompatibility complex. Wiley, New York

Klein J, O'hUigin C (1995) Class *II B Mhc* motifs in an evolutionary perspective. Immunol Rev 143:88–111

Klein J, Satta Y, O'hUigin C, Takahata N (1993) The molecular descent of the major histocompatibility complex. Annu Rev Immunol 11:269–295

Kreitman M, Hudson RR (1991) Inferring evolutionary histories of the *Adh* and *Adh-dup* loci in Drosophila *melanogaster* from patterns of polymorphism and divergence. Genetics 127:565–582

Li W-H, Graur D (1991) Fundamentals of molecular evolution. Sinauer Ass. Inc., Sunderland, Massachusetts

Li W-H, Wu C-I, Luo C-C (1984) Nonrandomness of point mutations as reflected in nucleotide substitutions in pseudogenes and its evolutionary implications. J Mol Evol 21:58–71

Lopez-Martinez N (1977) Revision sistematica biostratigraphica de los lagomorpha (mammalia) del neogeus y cuarternario de Espana. Doctoral Thesis, Madrid

Lummus Z, Cebra JJ, Mage R (1967) Correspondence of the relative cellular distribution and serum concentration of allelic allotypic markers in normal and allotype suppressed heterozygous rabbits. J Immunol 99:737–740

Mage RG (1967) Quantitative studies on the regulation of expression of genes for immunoglobulin allotypes in heterozygote rabbits. Cold Spring Harb Quant Biol 32:203–210

Mage RG (1981) The phenotypic expression of rabbit immunoglobulins: a model of the complex regulated gene expression and cellular differentiation. In: Inman FR, Mandy WJ (eds) Contemp Top Mol Immunol, Vol. 8. Plenum Press, New York, pp 89–112

Mage RG, McCartney-Francis NL, Kosmatsu M, Lamoyi E (1987) Evolution of genes for allelic and isotypic forms of immunoglobulin kappa chains and of the genes for T-cell receptor beta chains in rabbits. J Mol Evol 25:292–299

Mariamé B, Akimenko MA, Rougeon F (1987) Interallelic and intergenic conversion events could induce differential evolution of the two rabbit Ig kappa light chain genes. Nucleic Acids Res 15:6171–6179

McCartney-Francis N, Skurla RM Jr, Mage RG, Bernstein KE (1984) Kappa chain allotypes and isotypes in the rabbit: cDNA sequences of clones encoding b9 suggest an evolutionary pathway and possible role of the interdomain disulfide bond in quantitative allotype expression. Proc Natl Acad Sci USA 81:1794–1798

Merrit TJS, LaForest S, Prestwich GL, Quattro JM, Vogt R (1998) Patterns of gene duplication in Lepidopteran pheromone binding proteins J Mol Evol 46:272–276

Monnerot M, Vigne J-D, Biju-Duval C, Casane D, Callou C, Hardy C, Mougel F, Soriguer R, Dennebouy N, Monoulou J-C (1994) Rabbit and man: genetic and historical approach. Genet Select Evol 26:167s-182 s

Mougel F (1997) Variation de trois types de marqueurs génétiques dans l'évolution de l'espèce *Oryctolagus cuniculus*: aspects moléculaires et relations avec la biologie et la structure des populations. Doctoral thesis Université de Paris Sud, UFR scientifique d'Orsay, Orsay, France

Nei M (1973) Analysis of gene diversity in subdivided populations. Proc Natl Acad Sci USA 70:3321–3323

Nei M, Hughes AL: Polymorphism and evolution of the major histocompatibility complex loci in mammals. In: Selander R, Clark A (eds) (1994) Evolution at the molecular level. Whittam Sinauer, Sunderland, Massachusetts, pp 222–247

Ngai J, Dowling MM, Buck L, Axel R, Chess A (1993) The family of genes encoding odorant receptors in channel catfish. Cell 72:657–666

Oudin J (1960) Allotypy of rabbit serum proteins: I Immunochemical analysis leading to the individualization of seven main allotypes. J Exp Med 112:107–124

Reisfeld RA, Dray S, Nisonoff A (1965) Differences in amino acid composition of rabbit immunoglobulin light chains controlled by allelic genes. Immunochemistry 2:155–167

Short JA, Sethupathi P, Zhai SK, Knight KL (1991) *VDJ* genes in *VHa2* allotype-suppressed rabbits: limited germline *VH* gene usage and accumulation of somatic mutations in *D* regions. J Immunol 147:4014–4018

Sidney J, Grey HM, Kubo RT, Sette A (1996) Practical, biochemical and evolutionary implications of the discovery of *HLA* class I supermotifs. Immunol Today 17:261–266

Slierendregt BL, van Noort JT, Bakas R, Otting N, Jonker M, Bontrop RE (1992) Evolutionary stability of transspecies major histocompatibility complex class II *DRB* lineages in human and rhesus monkeys. Hum Immunol 35:29–39

Sogn JA, Kindt TJ (1976): A genetic polymorphism in the constant region of rabbit b4 kappa chain. J Exp Med 143:1475–1482

Takahata N, Nei M (1990) Allelic genealogy under overdominance and frequency-dependent selection and polymorphism of the major histocompatibility complex loci. Genetics 124:967–978

Takahata N, Satta Y, Klein J (1992) Polymorphism and balancing selection at major histocompatibility complex loci. Genetics 130:925–938

Tanaka T, Nei M (1989) Positive darwinian selection observed at the variable regions of immunoglobulins. Mol Biol Evol 4:447–459

Tonegawa S (1983) Somatic generation of antibody diversity. Nature 302:575–581

Vachot A-M (1996) Homologies et singularités interspecifques (*Ranidae*) et intraspecifiques (lapins). Doctoral Thesis, Université de Paris XI Orsay, Paris, France

Van der Loo W (1987) Studies on the adaptive significance of the immunoglobulin polymorphisms (Ig allotypes) in wild rabbit. In: Dubiski S (ed) The rabbit in contemporary immunological research. Longman Scientific & Technical, John Wiley and Sons, New York, pp 164–190

Van der Loo W (1993) Variance analysis of immunoglobulin alleles in natural populations of rabbit (*Oryctolagus cuniculus* L): The extensive interallelic divergence at the *b* locus could be the outcome of overdominant-type selection. Genetics 134:171–187

Van der Loo W, Arthur CP, Drees-Wallage M, Richardson B (1987) Nonrandom allele associations between unlinked protein loci: are the polymorphisms of the immunoglobulin constant regions adaptive? Proc Natl Acad Sci USA 84:3075–3079

Van der Loo W, Boussès P, Arthur CP, Chapuis J-L (1996) Compensatory aspects of allele diversity at immunoglobulin loci: gene correlations in rabbit populations devoid of light chain diversity (*Oryctolagus cuniculus* L.; Kerguelen Islands). Genetics 144:1181–1194

Van der Loo W, Bouton CE, Sanchez ME, Mougel F, Castién E, Hamers R, Monnerot M (1995) Characterization and DNA sequence of the b6w2 allotype of the rabbit immunoglobulin kappa 1 light chain (*b* locus). Immunogenetics 42:333–341

Van der Loo W, Brys L, Gobert M, Vanderveken M (1983) Mouse hybridoma lines specific for genetic markers of the rabbit immunoglobulins (allotypes). In: Peeters H (ed) Protides of the biological fluids (30th Colloquium), Pergamon Press Oxford, New York, pp 469–472

Van der Loo W, Ferrand N, Soriguer R (1991) Estimation of gene diversity at the *b* locus of the constant region of the immunoglobulin light chain in natural populations of European Rabbit (*Oryctolagus cuniculus*) in Portugal Andalusia and on the Azorean Islands. Genetics 127:789–799

Van der Loo W, Mougel F, Bouton CE, Sanchez ME, Monnerot M (1998) The allotypic patchwork pattern of the rabbit *IGKC1* allele *b5wf*: genic exchange or common ancestry? Immunogenetics 48:7–14

Van der Loo W, Verdoodt B (1992) Patterns of interallelic divergence at the rabbit *b* locus of the immunoglobulin light chain constant region are in agreement with population genetic evidence for overdominance-type selection. Genetics 132:1105–1117

Wayne ML, Contamine D, Kreitman M (1996) Molecular population genetics of *ref(2)P*, a locus which confers viral resistance in Drosophila. Mol Biol Evol 13:191–199

Webb NJ, Ibrahim KM, Bell DJ, Hewitt GM (1995) Natal dispersal and genetic structure in a population of the European wild rabbit (*Oryctolagus cuniculus*). Mol Ecol 4:239–247

Weir BS, Cockerham CC (1989) Complete characterisation of disequilibrium at two loci. In: Feldman ME (ed) Mathematical evolutionary theory. Princeton University Press, Princeton, New Jersey, pp 86–110